AD_____

GRANT NUMBER DAMD17-96-1-6084

TITLE: Metastic Regulation of Differential Splicing of CD44

PRINCIPAL INVESTIGATOR: Susan M. Berget, Ph.D.

CONTRACTING ORGANIZATION: Baylor College of Medicine
Houston, Texas 77030-3498

REPORT DATE: August 1998

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1998	Annual (15 Jul 97 - 14 Jul 98)

4. TITLE AND SUBTITLE Metastic Regulation of Differential Splicing on CD44

5. FUNDING NUMBERS
DAMD17-96-1-6084

6. AUTHOR(S)
Susan M. Berget, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
Houston, Texas 77030-3498

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

CD44 is a multi-functional adhesion molecule that can undergo alternative RNA splicing to generate multiple isoforms bearing different extracellular domains, a sub-set of which have been correlated with metastasis in breast cancer. Using a mouse model of mammary carcinogenesis we have observed that CD44 variable splicing alters during both normal breast development and mammary cancer. Pre-neoplasias, a tissue difficult to assay in humans exhibited alternative processing patterns similar to that in pregnant and lactating breast. The transition from pre-neoplasia to neoplasia was accompanied by strong increases in CD44 variable splicing. During the same progressions, alternative splicing factors expression changed. Co-transfection of some of these splicing factors along with a CD44 reported gene containing variable exons 4 and 5 demonstrated that CD44 splicing could be altered by these factors. Both positive and negative factors for CD44 splicing were identified. The concentration of both altered during neoplasia. These results suggest that alterations in splicing factors during early stages of breast cancer can have marked effects on RNA splicing to produce radically different proteins with different properties than those in normal cells.

14. SUBJECT TERMS Breast Cancer

15. NUMBER OF PAGES
40

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

SMB Where copyrighted material is quoted, permission has been obtained to use such material.

SMB Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

SMB In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

SMB In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

SMB In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

SMB In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Susan Berget 8/10/98

PI - Signature Date

(4) TABLE OF CONTENTS

(5) INTRODUCTION

Alternate pre-mRNA processing contributes significantly to the developmental regulation of gene expression in humans. It is estimated that as many as 25% of human genes utilize alternative RNA processing to produce subtly or grossly altered gene products (Moore *et al.*, 1993; Rio *et al.*, 1993; Berget *et al.*, 1995; Kramer *et al.*, 1996; Norton *et al.*, 1994)). During the last year, reports have emerged indicating that cancerogenesis induces changes in alternative processing (Lee *et al.*, 1997; Zhu *et al.*,1997; Silberstein *et al.*, 1997). Some of the genes targeted for these changes are receptor proteins suggesting that the induced alterations in splicing could have pronounced consequences for cellular behavior. Little, however is known about how RNA processing factors are altered during tumorigenesis so as to affect processing outcomes (Rio *et al.*, 1993; Kramer *et al.*, 1996; Norton *et al.*, 1994). The arginine-serine-rich (SR) proteins (Fig. 1a) constitute a family of splicing factors that recognize both splice sites and exonic splicing enhancers (Fu *et al.*, 1995; Zahler *et al.*, 1993; Valcarcel *et al.*, 1996), and influence alternative processing decisions when their relative concentrations are altered *in vivo* or *in vitro* (Ge *et al.*, 1990; Krainer *et al.*, 1990; Caceres *et al.*, 1994; Wang *et al.*, 1995). Furthermore, individual SR proteins have distinct tissue distributions (Screaton *et al.*, 1995; Zahler *et al.*, 1992). These observations have led to suggestions that alterations in the levels of SR proteins could be determinative for alternative splicing during development. This class of proteins, therefore, become attractive candidates for factors whose activity changes during tumorigenesis.

Two SR proteins, tra and tra-2 are important regulators of alternative splicing that are determinative for sexual development in *Drosophila*. These two proteins bind to CA-rich exon splicing enhancer elements within the alternatively-recognized female-specific exon within the

doublesex gene and cause exon recognition. Two alleles for tra-2, hTra-2α and hTra-2β have recently been observed in humans (Beil et al., 1997; Tacke et al., 1998). Although not yet linked to sexual development in human, the two proteins have very similar properties to the Drosophila homologues and can even replace functional tra-2 in fly sexual development. Similar tra-2 protein has also been found to be important for sexual development in C. elegans, raising the possibility that tra and tra-2 play important roles in sexual development in humans. Human tra-2 proteins have been observed to demonstrate preferential binding for A-rich sequences during an in vitro selection strategy (Tacke and Manley, 1998). Their natural targets, however, are unknown. Their connection with sexual development in other species, however, suggests that tra or tra-2 could play a role in breast development or tumorigenesis.

An ideal system in which to study the relationship between processing and the progression of cancer would be one in which both normal and abnormal development could be compared and studied in an isogenic and manipulatable background. Such a system is available in an established *in vivo* mouse model of mammary tumorigenesis (Kittrell *et al.*, 1992; Medina *et al.*, 1996) that provides access to normal mammary tissues from mature virgin, pregnant and lactating females, as well as an extensively characterized set of preneoplasias and their corresponding adenocarcinomas and metastases. In this model system, preneoplastic outgrowth lines are serially transplanted and maintained for extended periods of time (up to 12 months) in the mammary fat pads of syngenic female BALB/c mice. Each outgrowth line is characterized by a specific rate for development of adenocarcinomas and subsequent metastases (Kittrell *et al.*, 1992; Medina *et al.*, 1996). The outgrowths are clonal cell populations as determined by oncogene analysis (Cardiff *et al.*, 1988; Jerry *et al.*, 1993).

Using this system, we characterized SR protein expression during tumorigenesis. At the same time we examined the alternative splicing of a pre-mRNA, that coding for CD44, whose splicing had been reported to be determinative for metastasis (Gunthert *et al.*, 1991). CD44, is a widely expressed cell adhesion molecule and trans-membrane glycoprotein that mediates a variety of cell-cell and cell matrix interactions (Fox *et al.*, 1994; Mackay *et al.*, 1994; Stamenkovic *et al.*, 1989). A number of CD44 isoforms are created via alternative pre-mRNA splicing. Within the CD44 gene is an internal cassette of 10 alternatively spliced exons, all in the same translational reading frame, and coding for specific extracellular domains of the CD44 protein (Screaton *et al..*, 1992; Fig. 1b). Combinations of these variable exons lead to a variety of CD44 isoforms (Screaton *et al..*, 1992; Haynes *et al.*, 1990; Gunthert, 1993). Certain isoforms, especially those including variable exons v5, v6, and v7, have been implicated in the metastasis of several malignancies (East *et al.*, 1993; Cannistra *et al* ., 1995; Stickeler *et al.*, 1997; Wielenga *et al.*,1993) and correlated with survival in human breast cancer (Joensuu *et al.*. 1993; Kaufmann *et al.*, 1995). Recently it has been shown that alternative splicing of CD44 responds to signal transduction. Treatment of T-lymphoma cells that normally express a form of CD44 lacking variable exons, with phorbol esters, concanavalin A, or c-ras induces inclusion of variable exons (König et al, 1998). Thus, alternative splicing of CD44 may be involved in malignant transformation of tissues during tumorigenesis, and CD44 joins the collection of genes whose splicing is altered during cancer (Lee *et al.*, 1997; Zhu *et al.*,1997; Silberstein *et al.*, 1997).

Here we report that substantial alterations in SR splicing factor levels and CD44 alternative splicing accompanied the transition from preneoplasia to mammary adenocarcinoma, but not metastases. Furthermore, we observed that individual SR proteins responded differently

during development and neoplasia. Preneoplasias resembled pregnant mammary epithelia with respect to both SR expression and CD44 splicing patterns, although individual preneoplasias exhibited differences in their SR expression pattern. Alternative splicing of a control gene was not affected during preneoplasia or neoplasia, suggesting gene-specific alterations in RNA processing factors during early tumorigenesis, rather than a general increase in the generic processing machinery. In agreement with this interpretation, levels of constitutive splicing factors did not increase during neoplasia. Our results suggest that alterations in alternative processing may be important features of early and intermediate stages of mammary cancer.

(6) BODY

Experimental Methods, Assumptions and Procedures

Tissues. We used an *in vivo* mouse model of mammary development and tumorigenesis to study normal tissue, pregnant and lactating glands, preneoplastic lesions, 8 primary adenocarcinomas and organ metastases to the liver and lung. The TM preneoplastic outgrowth lines developed after transplantation of established mouse mammary epithelial cell lines (MMEL) into cleared mammary fat pads of 3 week old syngenic BALB/cMed mice (Kittrell *et al.*, 1992). The serially-transplanted outgrowths were removed either as preneoplasias at 8-12 weeks after transplantation or as tumors 5-7 months after transplantation. After removal, the preneoplastic outgrowths, primary adenocarcinomas, and metastases were frozen at -80 °C for further analysis. The biological properties of the outgrowth lines are shown in Table 1.

Epithelial cell isolation from mature virgin mammary tissue. Isolated mammary fat pads were minced and incubated in DMEM medium with 2 mg/ml collagenase A (Boehringer

Mannheim, Germany) and 100 U/ml hyaluronidase (Boehringer Mannheim, Germany) at 37 °C for 3 hours and slowly swirled. Afterwards the solution was centrifuged for 5 min at 1000 rpm. The supernatant, containing fat and single cells, was discarded and the pellet, containing mammary epithelial cells, was washed in PBS/5% FCS and stored at -80 °C for further analysis.

RNA Isolation, RT-PCR and Sequencing. Tissues were minced on dry ice in before they were treated with a tissue homogenizer (Polytron, Littau, Switzerland) in TRIzol solution (Gibco-BRL, Gaithersburg, MD) with 1ml per 100 mg of tissue for 3 times 10 sec. After 5 minutes of incubation at 26°C, 0.2 ml of chloroform per 1 ml of TRIzol were added and the solution was shaken vigorously, followed by incubation at 26°C for 5 minutes. After centrifugation for 15 min at 4°C and 12.000 x g the aqueous phase, containing the total RNA, was precipitated with 0.5 ml of isopropyl alcohol per 1 ml of TRIzol for 10 minutes at 26°C and centrifuged for 10 min at 12.000 x g at 4°C. The pellets were rinsed with 75% ethanol (1ml of ethanol per 1 ml of TRIzol reagent used for the initial homogenization), vortexed and centrifuged at 4°C for 5 min at 7.500 x g and finally dissolved in RNAse free water and stored at -80°C for further analysis.

Five micrograms of RNA were used for cDNA synthesis using M-MuLV reverse transcriptase (Perkin Elmer, Branchburg, NJ) and oligo-dT primers followed by PCR using primers specific for different SR proteins:

SRp20: 5' GCCGTGTAAGAGTGG, 3' AAGCTTCCTCCTTCTTGG;

ASF/SF2: 5' TCCGAGAACAGAGTGGTTGTCTCT, 3'GCTCTGGTAATTTACGGAAC;

SC35: 5' GACTCACAGCAGATCACGAGACC, 3' CCAGCCTTCCAGTCCCGCCTGTCAGAG;

SRp40: 5' TCGCAGACTACGGACCTCTC, 3' TGCA TCCCTTGGGTCCTCAAAT;

SRp55: 5' TCACGTAGTAGGTCACGGTCTC, 3' TTTCTCAGTAAATTTGGCCACA;

hTra-2α TATGATTACCGGTACAGAAGAAGG and GGAAACAAGCAACACAACTGG

hTra-2β TGACAATACATTCCACCACAG and GCATCAGCAATTTTTCTTCC

hnRNP A2: 5' ACAGTCTGTAAGCTTTCCCC, 3' CTGAAGCGACTGAGTCCGCG.

The standard form of CD44 mRNA (CD44 std) was detected using PCR primers for constitutive exons 5 and 16:

CD44 exon 16: CTCATAGGACCAGAAGTTGTGG

CD44 exon 5: ACCCCAGAAGGCTACATTTTGC

Inclusion of individual alternative exons was monitored using 3'primers specific for individual alternative exons and a primer complementary to constitutive exon 5:

CD44v3: (MASV3)CTTCATCATCATCAATGCCTGATCC,

CD44v5: (MASV5)TTGTA GCATGTGGGGTCTCCTCTT,

CD44v6: (MASV6) CCTTCTGTCACATGGGAGTCTTCA,

CD44v7: (MASV7) GATGACCTTGTCCCATTGGATGTG,

CD44v8: (MASV8) CACTGAAAGTGGTC CTGTCCTGTT.

To monitor transfection of the beta-globin gene modified to include CD44 variable exon 4 and 5, primers specific for exons1 and 2 were employed as follows.

Exon 1 AGACACCATGCATGGTGCACC

Exon 2 CCATAACAGCATCAGGAGTG

These primers did not amplify any signal from untransfected HeLa cells. All amplification reactions used an ATP-labeled 5' primer. The PCR conditions for all amplifications were as follows: 25-35 cycles (see Figure Legends) of 94°C for 1 min, 59°C for 1 min and 72°C for 1.5 min. Gel electrophoresis of 35 μl of the 100 μl PCR reaction was performed in 6% denaturing acrylamide gels at 30-40 mA. Expected amplification products are as follows: 142 bp (SRp20),

308 bp (ASF/SF 2), 300 bp (SC35), 222 bp (SRp40), 359 bp (SRp55), 134 and 170 bp (hnRNP A2 with or without the alternative exon), and 221 bp (CD44 standard lacking any alternative exon). The expected amplification bands for alternatively-spliced CD44 mRNA containing only one variable exon are 207 bp (v5), 209 bp (v6) and 208 bp (v7). pBR 322/Hpa II markers were used in all displayed gels for product size determination. Identified amplification products resulting from the inclusion of one or two variable CD44 exons were sequenced to verify identity. After cloning of CD44 PCR products into a pCR 2.1 vector (TA cloning kit ,Invitrogen, CA) sequencing of these clones was performed using the thermo sequenase method (Amersham, OH) and CD44 exon specific primers mentioned above.

Protein Isolation and Western Blot Analysis. Total cellular protein was isolated from the interphase and phenol phase from the initial homogenate after precipitation of the DNA with 0.3 ml 100% of ethanol per 1 ml of TRIzol used for the initial homogenization. The samples were stored at 26°C for 5 minutes and afterwards centrifuged at 2000 x g for 5 minutes at 4°C. The phenol-ethanol supernatant was than precipitated with 1.5 ml isopropyl alcohol per 1 ml TRIzol used for initial homogenization. After storage for 10 min at 26°C the protein precipitates were centrifuged at 12000 x g at 4°C for 10 minutes. Protein pellets were washed three times in 0.3 M guanidine hydrochloride in 95% ethanol for 20 minutes at 26°C and centrifuged at 7500 x g for 5 minutes at 4°C. The protein pellets were vortexed after the final wash in 2 ml 100% ethanol, stored for 20 minutes at 26°C and finally centrifuged at 7500 x g for 5 minutes at 4°C. The pellets were resuspended in 1 % SDS solution and incubated at 50°C for complete dissolution. Insoluble material was removed by centrifugation at 10000 x g for 10 minutes at 4°C. Supernatants were stored for further analysis at -80°C. Gel electrophoresis of 20 μg of total

protein was performed using a 10% SDS-Page gel at 100 volts. Afterwards the gels were electroblotted on a PVDF transfer membrane (PolyScreen, NEN Life Science, Boston, MA,) at 100 Volts for 2.5 hours at 4°C. After blocking in 5% Blotto/PBST membranes were incubated with mouse IgM mAb 104 (1:5 dilution), the SRp20 specific mouse IgG Ab (1:200 dilution) or polyclonal rabbit anti-U2AF Ab (1:6000 dilution) The Western blots were stained by chemiluminescence (NEN Life Science) using appropriate horseradish-peroxidase labeled anti IgM antibody for mAb 104 (Pierce, Rockford, IL), or anti IgG antibody for SRp20 Ab and U2AF Ab at a dilution of 1:3000.

Results and Discussion

Biological System. For this study, two mammary preneoplasias and a number of tumors were utilized. The tumorigenic and metastatic properties of the two preneoplastic outgrowth lines TM2L and TM40 are summarized in Table 1. Both TM2L and TM40 produce type B mammary adenocarcinomas by 12 months, at an incidence of 21% or 51%, respectively. The main difference between the two lines is the significant metastatic growth, primarily in lungs and occasionally in liver, of TM40.

Table 1 Tumorigenic characteristics of TM preneoplastic outgrowth lines

Outgrowth Line	Transplant Generation	Tumors/Transplants (%)	TE_{50} (months)[a]	% Metastases[b]
TM2L	21-28	21/98 (21)	9.5	0
TM40	4-13	80/157 (51)	11.0	73

[a] TE_{50} refers to time for 50% of the transplants to form tumors.

[b] Metastases were detected primarily in lung and secondarily in liver in mice 6-8 weeks after surgical removal of the primary tumors (11/15 animals). Metastases are rarely detected in animals bearing primary tumors.

Stepwise alterations in expression of SR proteins during development of mammary cancer. SR proteins are characterized by an extensive region containing arginine and serine (Fig. 1a) which is extensively phosphorylated and the epitope for a monoclonal antibody, mAb104 (Zahler *et al.*, 1993; Valcarcel *et al.*, 1996). To examine alterations in splicing factors during development of mammary cancer, we examined the expression of SR proteins in normal and neoplastic mammary tissues by Western Blots using mAb104. In HeLa cells, mAb104 recognizes approximately equal amounts of the major SR proteins SRp75, SRp55, SRp40, SRp30a, SRp30b, and SRp20 of molecular weights 75, 55, 40, 30 and 20 kDa, respectively (Zahler *et al.*, 1993). Mammary epithelial cells from pregnant or lactating animals expressed a sub-set of the major SR proteins, yielding dominant expression of SRp75 and one or more proteins of the SRp30 class (Fig. 2a, lanes 1, 2). One of the preneoplasias, TM2L, yielded an SR expression pattern similar to that of pregnant or lactating mammary gland (Fig. 2a, lane 7). The other tested preneoplasia, TM40, in addition, demonstrated detectable levels of SRp55 and SRp40, suggesting induction of synthesis of these SR proteins during the preneoplastic process (Fig. 2a, lane 8). Overall levels of the SR proteins were not dramatically changed in the TM40 preneoplasias. In contrast, all tested adenocarcinomas and their metastases were characterized by high level expression of the entire spectrum of SR proteins detected by mAb104 (Fig. 2a, lanes 4-6 and data not shown), indicating both a significant induction of SR proteins in neoplasias and a change in relative abundance of the members of the family compared to growing normal and preneoplastic mammary epithelia.

The utilized SR-specific antibody will not recognize SR proteins if they are not phosphorylated. To estimate if phosphorylation changes were responsible for altered levels of SR proteins, we repeated the analysis using a second antibody specific for one of the SR proteins

SRp20. This is antibody was raised against a SRp20 peptide from the region of the protein upstream of the phosphorylated SR region (Neugebauer and Roth, 1997). As shown in Figure 2b, SRp20 levels detected by this antibody resembled those detected by the mAb104. SRp20 levels increased in the neoplasias compared to virgin or pregnant gland and remained high in metastatic tissue. Because neoplasia is accompanied by increased growth, we also asked if generic splicing factors increased as a percentage of cellular proteins through our disease paradigm. Figure 2c shows a Western blot of normal and neoplastic tissues probed with an antibody specific for the constitutive splicing factor $U2AF^{65}$. As seen with this blot, U2AF levels remained a constant percentage of cellular proteins in these samples. Similar results were observed when the blots were probed with antibodies directed against small nuclear ribonucleoproteins (snRNPs) (data not shown). Therefore, the alterations in SR proteins occurring upon pre-neoplasia were not the result of a general induction of all factors required for splicing; instead only a sub-set of the major SR proteins were induced.

To confirm that the changes in SR protein levels observed in Figure 2 reflected altered mRNA synthesis rather than altered protein phosphorylation, we monitored SR protein mRNA levels using RT-PCR (Fig. 3). This approach was especially useful in examining expression of SRp20 that reacts relatively weakly with the mAb104 antibody (Fig. 3a). To control for non-linear estimations of mRNA levels using RT-PCR, SRp20 mRNA levels were also determined by RNase protection (data not shown). As observed by analysis with the Western blotting, tumorigenesis was accompanied by a major increase in the levels of SRp20 mRNAs, indicating that the increased levels of SRp20 in tumors seen with western blotting reflected increased mRNA synthesis rather than increased phosphorylation of an existing protein population (Fig.3a). RNase protection assays confirmed these findings with increased SRp20 mRNA levels

in tumorigenic tissues (data not shown). Similar results were observed when mRNA populations coding for ASF/SF2, SC35 (Fig.3b), SRp40, and SRp55 (Fig. 3c). All five tested SR mRNAs were significantly induced, suggesting that this family of proteins may be under similar transcriptional regulation.

RT-PCR revealed the presence of low levels of several of the tested mRNAs in preneoplasias and pregnant gland that had been difficult to assess by Western blotting. For example, SRp20 was detected in both lactating mammary tissue and in the preneoplasia TM40 by RT-PCR, even though neither had been easily visible by Western analysis, suggesting that low level synthesis of multiple SR proteins occurs in normal mammary gland epithelia. Levels of all tested SR mRNAs were noticeably higher in the TM40 preneoplasia than the TM2L neoplasia, suggesting important differences in these two lines of cells. Because most of the tested SR mRNAs were detectable albeit at a very low level in both preneoplasias by RT-PCR, this assay suggested that the difference in SR expression between the TM2L and TM40 preneoplasia is one of expression level.

Expression of SR proteins known to be involved in sex determination during mammary carcinogenisis. Two of the SR proteins for which excellent genetic evidence exists to document their importance in alternative splicing are transformer (tra) and transformer 2 (tra-2).

TISSUE/CELL TYPE	PERCENTAGE OF SAMPLES POSITIVE FOR TRA-2	
	hTRA-2α	**hTRA-2β**
Virgin gland	0% (3/3)	0% (3/3)
Pregnant gland	100% (2/2)	100% (2/2)
Lactating gland	100% (2/2)	100% (2/2)
Preneoplasias	40% (2/5)	50% (3/6)
Tumors	100% (5/5)	100% (5/5)

Table 2. Expression of tra-2 During Mammary Development and Tumorigenesis

Initial experiments with these proteins were done in Drosophila, but recent evidence indicates

that human tra and tra-2 may also play a role in sex determination in humans. We were especially interested in examining the expression of these proteins during mammary carcinogenesis because the sequences known to be recognized by these proteins are common exon enhancer sequences. In particular these sequences are frequently found in CD44, the gene we are interested in correlating with breast cancer and metastasis.

We examined the expression of tra-2 across our mouse model of mammary carcinogenesis. In humans there are two tra-2 genes, hTra-2α and hTra-2β that have distinguishable sequences that can be specifically amplified. In virgin gland we detected little expression of mRNA coding for either form of the protein (Table 2). Both mRNAs, however, were expressed during pregnancy and lactation. Therefore, tra-2 expression increased during mammary development. Preneoplasias varied as to the presence of tra-2 mRNAs. Of the 6 different neoplasias tested, approximately 50% expressed one or both. In contrast 100% of the tested tumors expressed both hTra-2α and hTra-2β. In addition the level of expression of tra-2 was higher in the tumors than in the preneoplasias. This analysis suggests that tra-2 responsive genes should switch splicing patterns during tumorigensis and these changes should mirror natural changes occurring during lactation.

The splicing of CD44 changes during progression of mammary cancer. Alteration in the relative levels of members of the SR proteins recognized by mAb 104 is known to affect splicing phenotypes. Thus, the changes in overall amount and, more importantly, in relative amounts of the SR proteins displayed in Figures 2 and 3 should be accompanied by changes in alternative processing of target genes. To examine this possibility we looked at the splicing of CD44, a gene with multiple alternative exons. Some of these exons have sequences similar to the purine-rich

splicing enhancers known to be the binding sites for the mAb 104-reactive SR proteins (Black, 1995; Reed, 1996). Splicing of CD44 was monitored by RT-PCR analysis of RNA from the same tissues used to examine SR proteins. CD44 mRNA is expressed as two basic forms, with or without one or more of the alternative internal exons. Expression of CD44 standard lacking all variable exons (CD44 std) was examined using primers complementary to constitutive exons 5 and 16, which border the central region of the CD44 gene encoding the ten alternative exons (Fig. 1b). RT-PCR of CD44 standard mRNA should produce an amplification band of 221 nucleotides; inclusion of one or more variable exons should produce larger bands. Because all of the variable exons are approximately the same size (Fig. 1b) inclusion of multiple variable exons produces an amplification pattern with a ladder of bands differing in length by the average variable exon length (approximately 115 nucleotides). This approach permits evaluation of the overall amount of CD44 mRNA present in examined tissues as well as an estimation of the extent of alternative splicing. Amplification revealed that mRNA encoding for CD44 std, lacking the variable exons, was expressed in all tissues examined (Fig. 4). All tested tumors and metastases, but not the precursor preneoplastic tissues, had increased CD44 mRNA, compared to virgin or pregnant gland. More interestingly, tumor RNA from adenocarcinomas or their metastases demonstrated noticeable alternative processing to reveal a family of amplification bands resulting from the inclusion of multiple variable exons. To exclude non-linear estimation of CD44 mRNA levels by relatively high amplification cycles, the amount of sample RNA's was titrated. The results showed linearity of RT-PCR signals in the range of used total RNA (data not shown).

To distinguish CD44-specific effects on alternative splicing associated with neoplastic transformation from generic effects on splicing activity associated with an accelerated growth

rate, we examined the amounts and spliced isoforms of an RNA that undergoes stochastic alternative splicing in many tissues. The hnRNP A2 gene is processed to produce two mRNA isoforms which alternatively include a 36 nucleotide variable exon. We observed relatively equal levels of total hnRNPA2 mRNA in all tissues examined and approximately equal relative levels of inclusion of the variable exon (Fig. 5). This result suggests that the changes in CD44 splicing occurring in mammary neoplasias is due to a change in the sub-set of splicing factors required to recognize the CD44 alternative exons and not a change in the generic splicing machinery.

CD44 metastatic exons are expressed at different levels during normal mammary gland development and tumorigenesis. Low-cycle RT-PCR with primers specific for constitutive exons 5 and 16 produces complicated amplification patterns because of the potential complexity of the family of CD44 mRNA sequences. To better examine variable exon splicing, RNAs from tissue samples were analyzed by RT-PCR using exon-specific primers for variable exons 5, 6, 7, or 8 and a consistent primer for constitutive exon 5 (Fig. 6, panels a-d), followed by cloning and subsequent sequencing of the PCR products This approach allowed examination of the inclusion of all alternative exons in the interval v1-v8. Both low and high cycle amplifications were performed. We concentrated on analysis of species containing variable exons v5, v6, or v7, because these exons have been associated with metastasis in other studies (Stickeler *et al.*, 1997; Heider *et al.*, 1995; Sinn *et al.*, 1995). Variable exon v8 was also of interest because it is one of three so-called "epithelial" exons, included in CD44 mRNA in a variety of epithelial cells.

Low cycle RT-PCR analysis of RNA from multiple tumors or their metastases using a primer specific for variable exon 5 (v5) indicated considerable production of RNAs containing

v5 with or without other upstream variable exons (Fig. 6a and data not shown). The inclusion patterns for all of the examined tumors and metastases were very similar, and indicated that most of the observed RNA included multiple variable exons from the v1-v5 region. In contrast, levels of inclusion of v5 were lower in normal dividing cells from pregnant or lactating mammary gland and in the two tested preneoplasias. Using 25-35 cycles of RT-PCR (Fig. 6a and b), the major amplified species in these tissues included either v5 alone or v5 and one other upstream variable exon. Cloning and sequencing of the larger amplification band indicated that 4/5 sequenced clones contained v4 in addition to v5. This species was more obvious in pregnant and lactating RNA than in RNA from the preneoplasias (Fig. 6a). We were unable to detect any RNA containing v5 in RNA isolated from mature virgin mammary gland (Fig. 6a, lane 3; and 6b, lane 2). Because of the abundance of fat tissue in mature virgin gland, we isolated epithelial cells from the extracted virgin mammary tissue. High cycle amplification of this RNA yielded no demonstrable inclusion of exon v5 either (data not shown). Therefore, for v5 splicing, three patterns of inclusion were observed by RT-PCR and sequencing: high level of inclusion of multiple exons along with v5 in tumor cells, low level of inclusion of v5 or v5+v4 in pregnant gland or preneoplasias, and no inclusion of v5 in normal virgin mammary gland. This observation suggests that preneoplasias adopt a splicing pattern similar to that induced in mammary epithelia during normal mammary development and that tumorigenesis is accompanied by induction of a pattern of frequent alternative splicing not found in normal development.

Analysis of alternative splicing of v6 or v7 gave results very similar to that discussed above for v5 with one exception (Fig. 6c and data not shown). Inclusion of both exons was included at high frequency along with other upstream exons in tumor cell RNA. RNA from

preneoplasias and pregnant gland demonstrated simpler RNAs containing fewer variable exons. For v6, these RNAs contained only v6; for v7, species containing v7 or v7+v4 were observed. In contrast to v5, both v6 and v7 were included in RNA from mature virgin gland epithelial cells (Fig. 6c, lane 4; and data not shown). Like v5, these observations suggest that splicing in preneoplasias is similar to splicing in normal proliferative mammary gland and that the conversion to neoplasia is accompanied by pronounced increases in mRNA including these variable exons. Therefore, for this system of mammary neoplasia, inclusion of v6 is not a marker for either tumorigenesis or metastasis. Instead, a better marker is increased inclusion of a number of CD44 variable exons including v6.

Analysis of RNA for the presence of v8 yielded a surprising result. Like the other variable exons, v8 was included in a number of RNA species in mammary tumor cells (Fig. 6d, lanes 5-6). Unlike v5-v7, however, no v8 inclusion could be observed in preneoplasias or RNA from developing mammary, despite their epithelial origin (Fig.6d, lanes 3-4). The observation of induction of v8 splicing in tumor RNA indicates that tumorigenesis is accompanied by alterations in RNA processing of CD44 that are never seen in normal development.

Splicing of CD44 variable exons 4 and 5 is altered by increased levels of SC35 and hTra-2β but not by SRp20, SRp40, or hTra-2α. The coordinate changes in SR protein expression and CD44 splicing during mammary development and carcinogenesis suggested that alterations in SR expression could be causative for changes in CD44 splicing. To address this question we embarked on a set of experiments to use transfection of cultured cells to monitor CD44 splicing under conditions in which SR protein level is modulated. For this purpose we created a mini-gene that includes two CD44 variable exons between two constitutive exons from the beta globin

gene (Fig. 7a). This mini-gene contains natural sequences from human CD44 from within the intron upstream of variable exon 4 down to sequences within the intron downstream of variable exon 5. This region included natural exons 4 and 5 and the intron between them. Transfected into HeLa cells this construct produced a mixed splicing pattern (Fig. 7b). The majority of the RNA produced (82%) resulted from skipping of both variable exons, reflecting the weakness of these variable exons. Approximately 15% of the RNA included both variable exons and a small percentage (3%) of the RNA included either variable exon 4 or variable exon 5, but not both exons.

We then co-transfected cDNA plasmids expressing one of the SR proteins. As can be seen in Figure 7b, transfection with SRp20 or hTra-2α had minimal impact on the splicing patterns. Transfection of SC35 strongly suppressed inclusion of either of the variable exons, suggesting a negative effect on exon recognition. Inclusion of both exons was reduced 10-fold, such that 1.5% of the RNA contained both exons, 4.5% contained one of the variable exons, and 94% contained neither. hTra-2β had an interesting property in that it increased the inclusion of single exons at the expense of both exons. Single exon inclusion rose to 8%.

Discussion. During the last several years, studies have suggested changes in pre-mRNA splicing in human malignancies (Lee *et al.*, 1997; Zhu *et al.*,1997; Silberstein *et al.*, 1997). Using an *in vivo* model of mammary development and tumorigenesis, we were able to study these questions with a focus on comparing splicing phenotypes and factors during normal mammary development to those occurring in development of mammary cancer. Neoplasia was accompanied by a dramatic increase in expression of the SR family of splicing factors recognized by mAb104, resulting in both an alteration in total abundance of these important

splicing factors and the relative abundance of the individual proteins. Given the number of genes whose splicing has been observed to be responsive to relative SR protein levels (Ge *et al.*, 1990; Krainer *et al.*, 1990; Caceres *et al.*, 1994; Wang *et al.*, 1995), this observation suggests that pronounced changes in alternative splicing of a number of pre-mRNAs should accompany mammary tumorigenesis.

The utilized model system also permitted examination of SR protein expression in proliferative epithelial cells from pregnant mammary gland as compared to those in preneoplasias. Mammary epithelia from pregnant mice expressed low amounts SR proteins with preferential expression of two members of the family of SR proteins recognized by mAb 104. The two preneoplasias characterized in this study, TM2L and TM40, differed in their SR protein expression pattern with a noticeably higher and more complex expression pattern in TM40 than TM2L. These two preneoplasias arose independently and have different, albeit modest, tumor producing capabilities. Both outgrowth lines contain wild-type p53 as determined by immunohistochemical staining and by nucleotide sequence analysis (Medina, unpublished). Both preneoplastic lines are ovarian hormone-independent and give rise to adenocarcinomas of type B. The main biological factors which serve to distinguish the two lines are the higher metastatic potential of TM40 and the histomorphology of the TM40 outgrowth (a ductule hyperplasia versus the more typical lobulo-alveolar hyperplasia exemplified by TM2L). It is difficult to correlate the differences in biological properties of the two preneoplasias with the expression of SR proteins. Alterations in relative SR protein levels, however, would be predicted to have pronounced effects on the constellation of proteins produced in a cell.

Analysis of constitutive splicing factors and a stoichastically alternatively processed gene indicated that the changes in SR expression we observed were not the result of induction of all

splicing during neoplasia. Levels of the constitutive factors U2AF^{65} and U snRNPs did not alter during preneoplasia and neoplasia as a percentage of cellular proteins. In addition, splicing of an hnRNP gene that is stoichastically alternatively spliced in a large number of non-neoplastic tissues was not altered in the investigated paradigm. Therefore, the alterations we observed occurred in a family of factors associated with alternative splicing only.

We also used this model system to investigate potential changes in alternative splicing in a gene whose splicing changes during cancer have been correlated to metastatic potential in human cancers and which has recently been shown to alter its splicing patter in response to signal transduction (König et al., 1998). As with SR proteins, striking increases in alternative splicing of CD44 were observed in adenocarcinomas as compared to either preneoplasias or proliferative mammary epithelia from pregnant animals. All of the tumors examined displayed a similar pattern of exon inclusion, in which multiple alternative exons were included in mRNA. Included in this set of exons were exons not normally present in epithelia of pregnant mammary gland or in preneoplasias. The splicing patterns of the adenocarcinomas was similar to that of the lung and liver metastases of these tumors, indicating few changes accompanied the metastases. In contrast to CD44 alternative splicing, the splicing of a control RNA did not change during the transition from preneoplasia to neoplasia. This difference suggests that the pronounced changes in splicing of CD44 observed during development of neoplasia arise from the induction of specialized splicing factors, not an increase in the generic splicing machinery.

When we examined CD44 variable exon splicing in epithelia from pregnant mammary gland, we observed low levels of inclusion of some but not all the variable exons. In particular, v6 and v7 were present in a sub-population of CD44 mRNA from both pregnant tissue and preneoplasias, indicating that these so-called "metastatic" exons are included in non-cancerous

mammary epithelial cells. In contrast, one variable exons, v5, was not expressed in virgin mammary gland, indicating that alterations in CD44 alternative splicing can occur during preneoplasia and resemble those occurring during pregnancy. Other variable exons that are not included in pregnant tissue (such as the "epithelial" exon v8) are also not expressed during preneoplasia, but are included in tumor RNA. Therefore, our analysis suggests that preneoplasias demonstrate a splicing pattern seen during normal breast development. Known to participate in cell-cell interactions and cell-matrix interactions, CD44 isoforms could be involved in the ductal outgrowth occurring during early pregnancy.

Inclusion of three of the CD44 variable exons (v5, v6, and v7) have been postulated to be markers for metastasis potential in several human cancers, including breast cancer. The initial interest in these exons arose from an elegant experiment demonstrating that transfection of a CD44 cDNA including exon v6 into a tumor cell of no metastatic potential increased the rate of metastasis when these cells were placed back into syngenic animals. Such experiments expressed the variant CD44 mRNA at high levels. Our analysis would suggest that the major change in CD44 splicing occurs during development of adenocarcinomas , not their metastases and that it may be the over-expression of variant CD44 mRNAs during the neoplastic state which is important for the metastatic process, rather then the production of any mRNA including v6. This difference may suggest why experiments monitoring inclusion of CD44 variable exons have failed to demonstrate a consistent correlation to disease progression in several clinical studies (Friedrichs *et al.*, 1995; Muller *et al.*, 1997; Tran *et al.*, 1997).

The recent report of a coupling between signal transduction and alternative splicing of CD44 concentrated on variable exon v5. Inclusion of this exon showed a pronounced response to induction by c-ras, TPA, or phorbol esters (König *et al.*, 1998). Exon v5 contains internal

sequences that resemble known purine-rich exon enhancers that bind SR proteins. These observations suggest that signal transduction cascades can lead to an induction of SR proteins needed for CD44 variable splicing. In light of these findings, our observation that inclusion of v5 is pronouncedly changed upon preneoplasia supports our connection between SR protein levels and CD44 alternative splicing.

Besides CD44 a number of other genes show altered RNA processing during tumorigenesis. A number of cellular receptors and hormone genes undergo alternative processing and this processing appears to alter during cancer. Because the changes in these proteins caused by alternative processing can have a pronounced effect on cellular function, it is important to understand the role of alternative processing in cancer and the mechanisms involved. Determination of splicing factor constellations patterns in preneoplastic lesions of the mammary gland could be very helpful to identify patients in high-risk situations to develop invasive breast cancer and subsequent organ metastasis. Distinct expression patterns of splicing factors could serve as new markers for metastasis in breast cancer, and splicing factors may represent targets for intervention in a subgroup of patients abnormally expressing certain factors.

Recommendations in Relation to the Statement of Work.

Our recent observations that we can make a mini-gene including CD44 variable exons 4 and 5 which responds to exogenously added splicing factors is very encouraging. It strongly indicates that we will be able to identify factors that are responsible for altering splicing in breast cancer.

We have elected to begin our study of individual factors and CD44 splicing with variable exon 4 and 5. These are attractive targets for several reasons. First, in our mammary tumor samples, especially the preneoplasias, exon 5 inclusion was accompanied by exon 4 inclusion suggesting that these exons are recognized as a pair. Secondly, the intron that separates this pair of exons is frequently retained in CD44 mRNA and inclusion of this intron has received attention as a marker for certain cancers. Third, there are sequences repeated within exons v4 and v5 and the intron separating them that are visually recognizable as binding sites for alternative splicing factors, suggested concerted recognition of the entire region. Fourth, transfection with tra-2, a known splicing regulator that binds to this recognizable sequence in other genes, has an effect on CD44 variable splicing. But most importantly, the inclusion of variable exon 5 has recently been reported to respond to signal transduction pathways (König et al., 1998), indicating important biological regulation of the inclusion of this exon.

During the next year we plan to further characterize the proteins that bind variable exon 4 and 5 as well as the intron between them.. We are testing a number of candidate proteins not mentioned in this report and have tantalizing first glimpses that there are additional important regulatory proteins that are important for this recognition event. We will analyze the requirement for these proteins for inclusion both *in vivo* and in our *in vitro* splicing extract. This combination of *in vitro* and *in vivo* approaches should permit careful dissection of the protein factors involved. We will of course also monitor the expression and activity of identified factors in our mammary model of breast cancer development and disease.

(7) CONCLUSIONS

Using a mouse model of mammary gland development and tumorigenesis we examined changes in both alternative splicing and splicing factors in multiple stages of mammary cancer. The emphasis was on the SR family of splicing factors known to influence alternative splicing in a wide variety of genes, and on alternative splicing of the pre-mRNA encoding CD44, for which alternative splicing has been implicated as important in a number of human cancers, including breast cancer. We observed step-wise and varied increases in expression of individual SR proteins and alternative splicing of CD44 mRNA during mammary gland tumorigenesis. Pronounced differences in both were observed during progression from preneoplasia to neoplasia. In contrast, little difference was observed between neoplasias and their metastases. Important changes in splicing and splicing factors also occurred early during preneoplasia and some of these changes mirrored or extended changes occurring during pregnancy. Given the ability of SR proteins to affect many alternative processing decisions, these data suggest that a number of pre-mRNAs may undergo changes in alternative splicing during the early and intermediate stages of mammary cancer.

(8) REFERENCES

Beil, B., G. Screaton, and S. Stamm. 1997. DNA Cell. Biol. **16:** 679-690.

Berget S. M. (1995). J. Biol. Chem., **270**, 2411-2414.

Black D. L. (1995). RNA, **1**, 763-771.

Caceres J. F, Stamm S, Helfman D. M. and Krainer A. R. (1994). Science, **265**, 1706-1709.

Cannistra S., Abu-Jawdeh G. and Niloff J. (1995). J. Clin. Oncol., **13**, 1912-1921.

Cardiff R. D. (1988). Anticancer Res., **8(5A)**, 925-933.

East J. A. and Hart I. R. (1993). Eur. J. Cancer, **29A(14)**, 1921-1922.

Fox S. B., Fawcett J., Jackson D. G , Collins I., Gatter K. C., Harris A. L., Gearing A. and Simmons D. (1993). Cancer Res., **54**, 4539-4546.

Friedrichs, K., Kugler, G., Franke, F., Terpe, H. J., Arlt, J., Regidor, P. A. and Gunthert, U. (1995). Lancet, **345**, 1237.

Fu, X-D. (1995). RNA, **1**, 663-680.

Ge. H, and Manley. J. L. (1990). Cell, **62(1)**, 25-34.

Gunthert, U., Hoffmann, M., Rudy, W., Reber, S., Zoller, M., Haussmann, I., Matzku, S., Wenzel, A., Ponta, H. and Herrlich, P. (1991). Cell, **65**, 13-24.

Gunthert, U. (1993). Curr. Microbiol. and Immunol., **184**, 47-63.

Haynes, B. F., Hua-Xin, L. and Patton, K. L. (1990). Cancer Cells, **3**, 347-350.

Heider, K. H., Mulder, J. W., Ostermann, E., Susanni, S., Patzelt, E., Pals, S. T. and Adolf, G. R. (1995). Eur J Cancer, **31A**, 2385-2391.

Jerry, D.J., Ozbun, M. A., Kittrell, F. S., Lane, D. P., Medina, D. and Butel, J. S. (1993). Cancer Res., **53(14)**, 3374-3381.

Joensuu, H., Klemi, P., Toikkanen, S. and Jalkanen, S. (1993). Am. J. Path., **143**, 867-874.

Kittrell, F. S., Oborn, C. J. and Medina, D. (1992). Cancer Res., **52**, 1924-1932.

Krainer, A. R., Conway, G. C. and Kozak, D. (1990). Cell, **62(1)**, 35-42.

Kramer, A. (1996). Annu. Rev. Biochem. , **65**, 367-409.

Kaufmann, M., Heider, K. H., Sinn, H. P., von Minckwitz, G., Ponta, H. and Herrlich, P. (1995). Lancet, **345**, 615-619.

König, H., Ponta, H., and Herrlich, P. (1998) EMBO J. **17,** 2904-2913.

Lee, M. P. and Feinberg, A. P. (1997). Cancer Res., **57**, 3131-3134.

Mackay, C., Terpe, H. J., Stauder, R., Marston, W. L., Stark, H. and Gunthert, U. (1994). J. Cell Biol., **124**, 71-82.

Medina, D. (1996). Mammary Tumor Cell Cycle, Differentiation and Metastases. Dickson, R. B. and Lippman, M. E. (eds) , Kluwer Academic Publishers: Norwell, pp.37-69.

Moore, M. J., Query, C. C. and Sharp, P. A. (1993). The RNA World. R. F. Gesteland and J. F. Atkins (eds.). Cold Spring Harbor Laboratory Press: New York, pp. 303-357.

Muller, W , Schneiders, A., Heider, K. H., Meier, S., Hommel, G. and Gabbert, H. E. (1997). J. Pathol., **183(2)**, 222-227.

Neugebauer, K. M. and Roth, .M. (1997a). Genes Dev. **11**, 1148-1159.

Norton, P. A. (1994). J. Cell Sci., **107**, 1-7.

Reed, R. (1996). Curr. Opin. Genet. Dev., **6**, 215-220.

Rio, D. C. (1993). Curr. Opin. Genet. Dev., **3**, 574-584.

Screaton, G. R., Bell, M. V., Jackson, D. G., Cornelis, F. B., Gerth, U. and Bell, J. I. (1992). Proc. Natl. Acad. Sci USA, **89**, 12160-12164.

Screaton, G. R., Caceres, J. F., Mayeda, A., Bell, M. V., Plebanski, M., Jackson, D. G., Bell, J. I. and Krainer, A. R. (1995). EMBO J., **14**, 4336-4349.

Silberstein, G. B., Van Horn, K., Strickland, P., Roberts, Jr. C. T. and Daniel, C. W. (1997). Proc. Natl. Acad. Sci. USA, **94**, 8132-8137.

Sinn, H. P., Heider, K. H., Skroch-Angel, P., von Minckwitz, G., Kaufmann, M., Herrlich, P. and Ponta, P. (1995). Breast Cancer Res Treat, **36(3)**, 307 313.

Stamenkovic, I., Amiot, M., Pesando, J. M. and Seed, B. (1989). Cell, **56**, 1057-1062.

Stickeler, E., Runnebaum, I. B., Moebus, V. J., Kieback, D. G. and Kreienberg, R. (1997). Anticancer Res., **17**, 1871-1876.

Tacke, R., M. Tohyama, S. Ogawa, and J.L. Manley. 1998. Cell **93:**139-148.

Tran, T. A., Kallakury, B. V., Sheehan, C. E. and Ross, J. S. (1997). Hum. Pathol., **28(7)**, 809-814.

Valcarcel, J. and Green, M. R. (1996). Trends Biochem. Sci., **21**, 296-301.

Wang, J. and Manley, J. L. (1995). RNA, **1(3)**, 335-346.

Wielenga, V. J., Heider, K. H., Offerhaus, G. J., Adolf, G. R., van den Berg, F. M., Ponta, H., Herrlich, P. and Pals, S. T. (1993). Cancer Res., **53**, 4754-4756.

Zahler. A. M., Lane, W. S., Stolk, J. A. and Roth, M. B. (1992). Genes Dev., **6**, 837-847.

Zahler, A. M., Neugebauer, K. M., Stolk, J. A. and Roth, M. B. (1993). Mol.Cell Biol. , **13(7)**, 4023-4028.

Zhu, X., Daffada, A. A., Chan, C. M. and Dowsett, M. (1997). Int. J. Cancer., 72, 574-580.

(9) APPENDICES

Figure Legends

Figure 1 SR proteins and CD44 architecture. (a) Schematic of a generic SR protein with two N-terminal RRM motifs that bind RNA (boxes) and extensive C-terminal RS (arginine-serine) domains (vertical lines). Thickness of vertical lines represents numbers of consecutive RS dipeptides. **(b)** CD44 exon structure. Constitutive exons 1-5 and 16-20 flank a central region encoding at least 10 alternatively included (variable) exons (v1-v10). The size of each variable exon is indicated. Arrows indicate primers used for RT-PCR analysis.

Figure 2 The expression of SR proteins is altered in mammary adenocarcinomas and in some preneoplasias. (a) Western Blot analysis of SR protein expression in different breast tissues using an antibody specific for a phosphorylated SR epitope (mAb104). Lane 1, 5-day pregnant gland; lane 2, 15-day pregnant gland; lane 3, 1-day lactating gland; lanes 4-6, independent adenocarcinomas of the breast; lane 7, pre-neoplasia TM2L; lane 8, preneoplasia TM40. Equal amounts of protein (20μg) were used for each lane. For the pregnant and lactating breast, and the preneoplasias, estimated epithelial cell content of the excised tissue exceeded 85%. **(b)** Western blot analysis using an antibody specific for SRp20. The utilized antibody (Neugebauer et al, 1997) is an anti-peptide antibody specific for the region of SRp20 between the RNA binding domain and the SR domain. Lane 1, mature gland; lane 2, preneoplasia TM2L; lane 3, preneoplasia TM40; lane 4, adenocarcinoma of the breast; lane 5, liver metastasis. **(c)** Western Blot analysis of constitutive splicing factor U2AF using a polyclonalclonal rabbit antibody raised against the the 65 kDA subunit without the SR domain.. Lane 1, preneoplasia TM2L; lane 2, preneoplasia TM40; lane 3, adenocarcinoma of the breast; lane 4, liver metastasis.

Figure 3. Alterations in levels of mRNAs coding for splicing factors during tumorigenesis. Levels of SRp20 revealed by RT/CRP (a). RT-PCR analysis of mRNAs coding for individual S/R proteins in breast tissues using primers specific for ASF/SF2, SC35 (b), SRp40 or SRp55 (c). **(a)** SRp20: Lane 1, 1-day lactating gland; lane 2, mature virgin gland; lane 3, preneoplasia TM2L; lane 4, preneoplasia TM40; lanes 5,-10 different adenocarcinomas of the mammary gland. **(b)** ASF/SF2 and SC35: Lane 1, preneoplasia TM40; lane 2, preneoplasia TM2L; lanes 3-6, adenocarcinomas of the mammary gland. **(c)** SRp40 and SRp55: Lane 1, preneoplasia TM40; lane 2, preneoplasia TM2L; lanes 3-5, adenocarcinomas of the mammary gland. Equal amounts (5μg) of total cell RNA were used for RT/PCR amplification using gene-specific primers. See Figure 5 for RT/PCR analysis of a control RNA in these RNA samples.

Figure 4. Alternative splicing of CD44 changes during mammary development and neoplasia. RT-PCR analysis of CD44 mRNA isoforms in breast tissues using primers specific for constitutive exons 5 and constitutive exon 16. Amplification was for 25 cycles. Boxes symbolize included exons in the amplification bands. Exon structures of amplifications products are indicated. Hatching indicates mixtures of included variable exons. Majority products are indicated where they existed. Lane 1, lactating gland; lane 2, pregnant gland; lane 3, mature virgin gland; lane 4, preneoplasia TM2L; lane 5, preneoplasia TM40; lanes 6-10, different adenocarcinomas. All lanes used equal amounts (5μg) of total cell RNA for amplification. See Figure 5 for RT/PCR analysis of a control RNA in these RNA samples.

Figure 5. Alternative splicing of mRNA coding for hnRNP A2 does not change through mammary development or neoplasia. Low cycle RT-PCR of hnRNP A2 mRNA in mammary

tissues using primers specific for two constitutive exons (white boxes) flanking a 36 nucleotide alternatively included exon (hatched box). Exon skipping produces a 135 nt band (bottom), inclusion in a 171 nt amplification product (top). Lane 1, 1 day lactating gland; lane 2, preneoplasia TM40; lanes 3-6 adenocarcinomas; lane 7, liver metastasis. Equal amounts of total cell RNA (5μg) were analyzed.

Figure 6. CD44 metastatic exons are expressed in normal development and tumorigenesis. RT-PCR analysis of CD44 mRNA variable isoforms using primers specific for constitutive exons 5 and variable exon v5 (a and b), variable exon v6 (c) or variable exon v8 (d). Amplification was for 25 (a,c,d) or 35 (b) cycles. Boxes symbolize included exons in the amplification bands. Exon structures of amplifications products are indicated. Hatching indicates mixtures of included variable exons. Majority products are indicated where they existed. **(a,c)** Lane 1, lactating gland; lane 2, pregnant gland; lane 3, mature virgin gland; lane 4, preneoplasia TM2L; lane 5, preneoplasia TM40; lanes 6-10, different adenocarcinomas. **(b)** Lane 1, pregnant gland; lane 2, mature virgin gland; lane 3, pre-neoplasia TM2L; lane 4, pre-neoplasia TM40. **(d)** Lane 1, lactating gland; lane 2, mature virgin gland; lane 3, pre-neoplasia TM2L; lane 4, preneoplasia TM40; lanes 5-6, adenocarcinomas. All lanes used equal amounts (5μg) of total cell RNA for amplification.

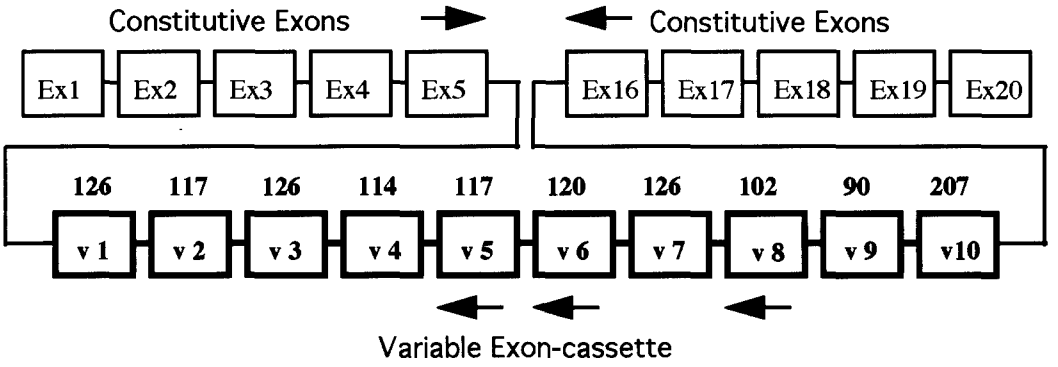
Figure 7. Over-expression of SC35 or hTra-2β alters CD44 splicing. a) Mini-gene to study CD44 variable exon splicing. A region of genomic CD44 DNA encompassing the last half of intron v3, exon v4, intron v4, exon v5, and half of intron v5 was introduced into a beta-globin expression vector lacking its middle exon. Exon and intron sizes are as indicated. b) The

construct in a was transfected into HeLa cells along with expression vectors for the indicated SR proteins. RNA phenotypes were assayed by 20 cycle RT/PCR using primers specific for vector sequences in exons 1 and 4. The various spliced product RNAs are indicated.

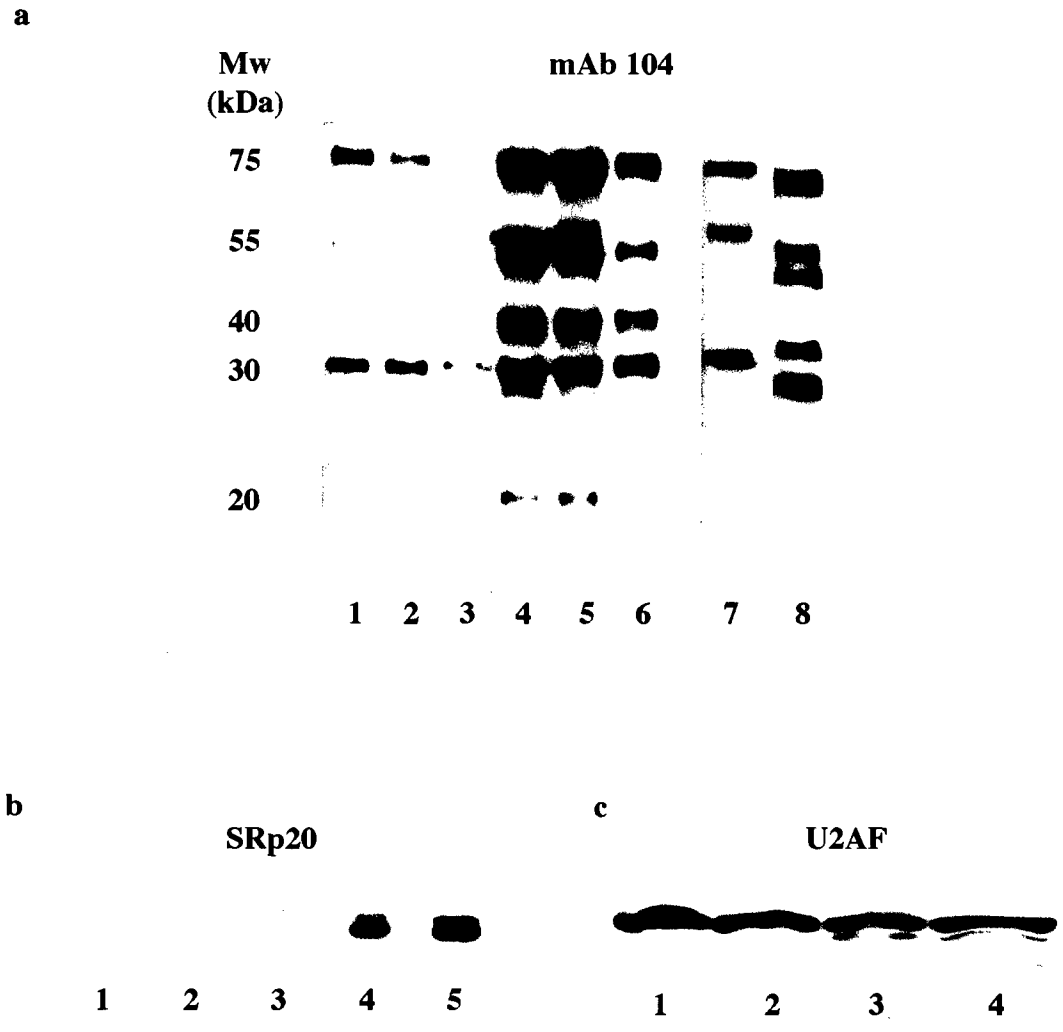
Stickeler et al., Fig. 1

a

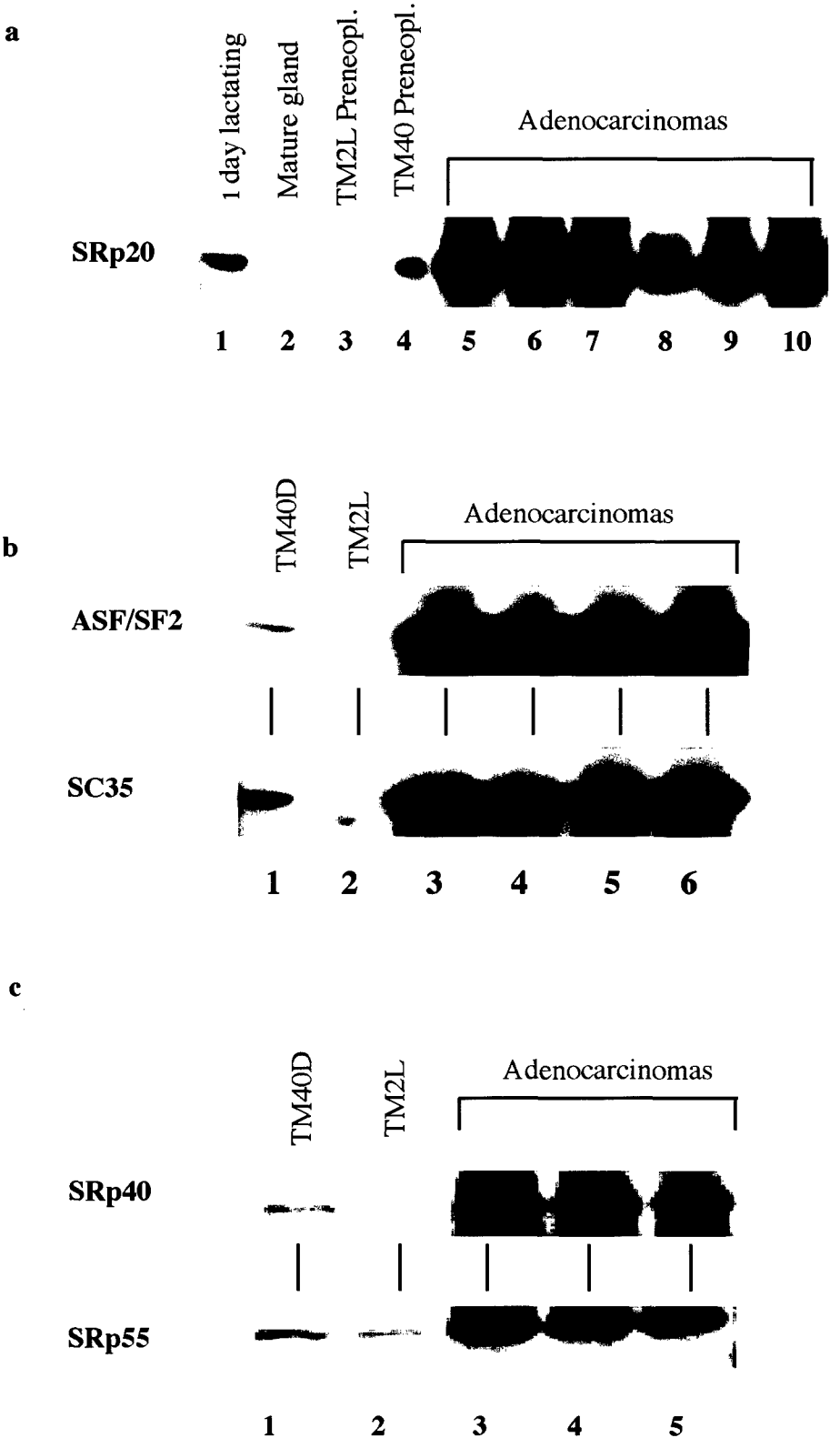
b



Stickeler et al., Figure 2



Stickeler et al., Figure 3



Stickeler et al., Figures 4 and 5

Fig. 4

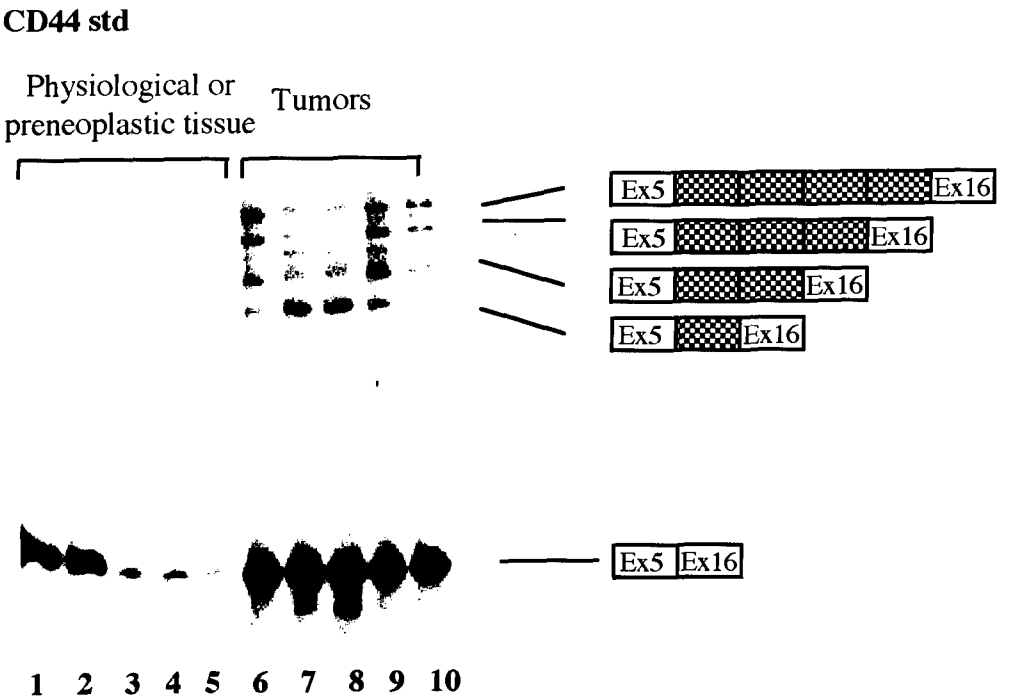
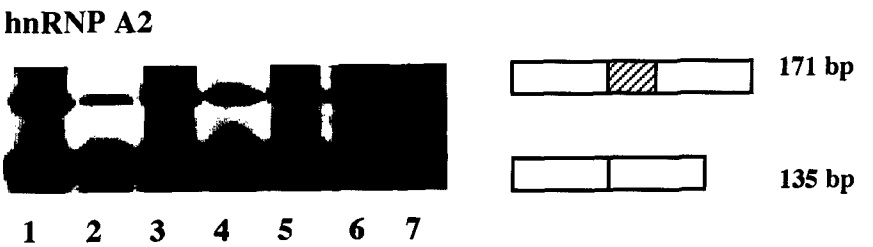
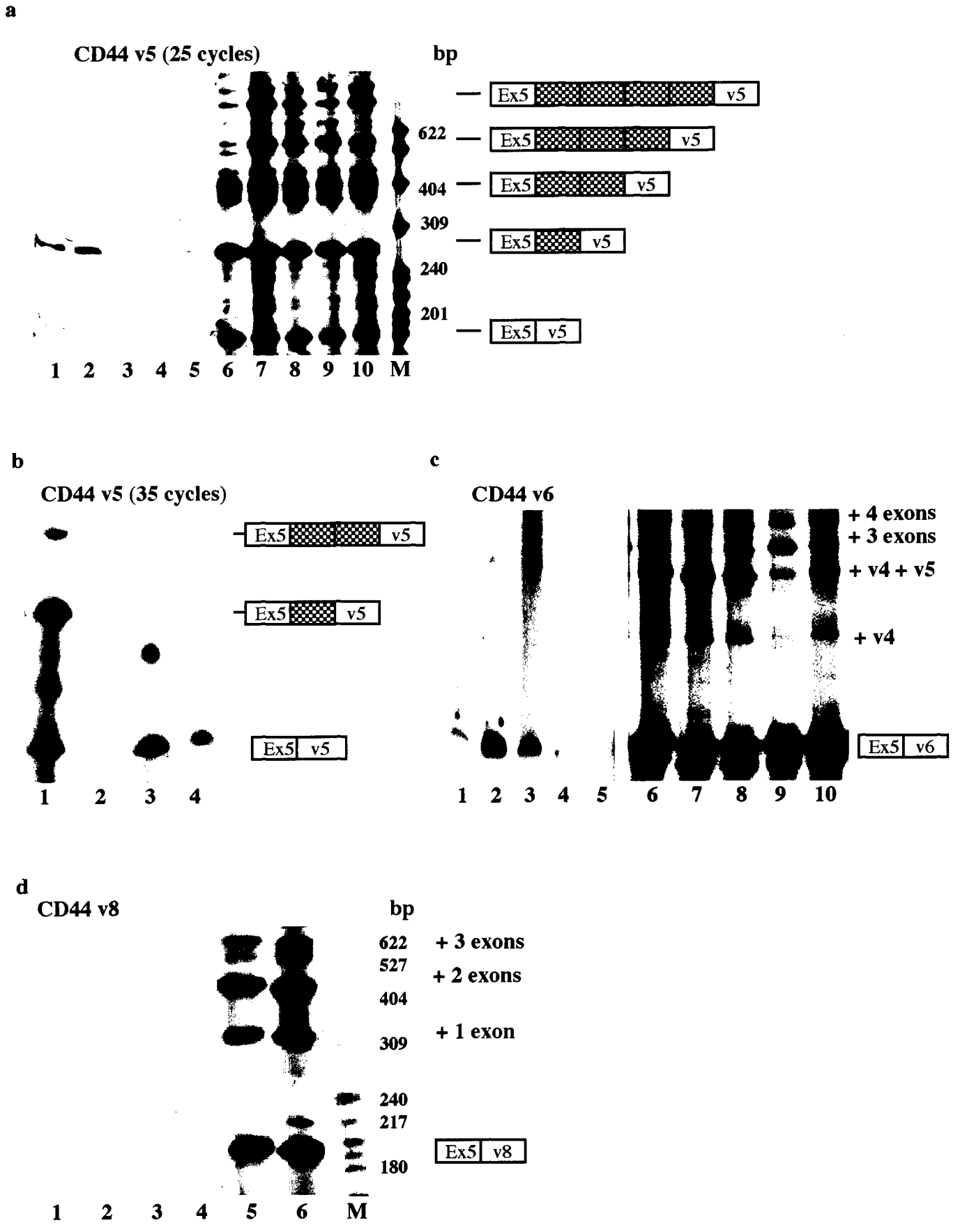


Fig. 5

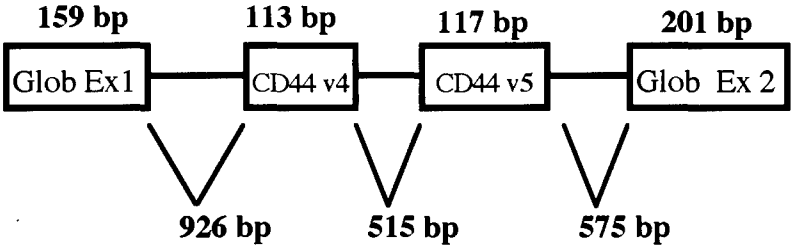



a
CD44 v5 (25 cycles)
bp
622
404
309
240
201
Ex5
v5
1 2 3 4 5 6 7 8 9 10 M
b
CD44 v5 (35 cycles)
Ex5
v5
1 2 3 4
c
CD44 v6
+ 4 exons
+ 3 exons
+ v4 + v5
+ v4
Ex5 v6
1 2 3 4 5 6 7 8 9 10
d
CD44 v8
bp
622
527
404
309
240
217
180
+ 3 exons
+ 2 exons
+ 1 exon
Ex5 v8
1 2 3 4 5 6 M

Stickeler et al., Figure 7

a

EST GLvII Construct



b

HeLa Transfection: ESTGlvII + Splicing Factors